

# Some properties of the stationary distributions of subcritical multitype Galton–Watson processes

Gábor Szűcs

Bolyai Institute, University of Szeged

## 1 Introduction

In our paper  $\mathbb{Z}_+$  denotes the set of non-negative integers. For any  $\mathbf{x} = (x_1, \dots, x_p)^\top \in \mathbb{R}^p$  and  $\mathbf{y} = (y_1, \dots, y_p) \in \mathbb{R}^p$  the notation  $\mathbf{x} \leq \mathbf{y}$  is understood componentwise, that is,  $\mathbf{x} \leq \mathbf{y}$  if and only if  $x_i \leq y_i$  for  $i = 1, \dots, p$ . The norm of the arbitrary vector  $\mathbf{x}$  is defined as  $\|\mathbf{x}\| = |x_1| + \dots + |x_p|$ . Also, using the Kronecker delta symbol  $\delta_{i,j}$  the system  $\mathbf{e}_i = (\delta_{i,1}, \dots, \delta_{i,p})^\top$ ,  $i = 1, \dots, p$ , stands for the usual base of the vector space  $\mathbb{R}^p$ .

## 2 Main results

The  $p$ -type Galton–Watson process  $\mathbf{X}_n = (X_{n,1}, \dots, X_{n,p})^\top$ ,  $n = 0, 1, \dots$ , is a sequence of  $\mathbb{Z}_+^p$  valued random vectors defined by the recursion

$$\mathbf{X}_n = \sum_{k=1}^{X_{n-1,1}} \boldsymbol{\xi}_1(n, k) + \dots + \sum_{k=1}^{X_{n-1,p}} \boldsymbol{\xi}_p(n, k) + \boldsymbol{\eta}(n), \quad n = 1, 2, \dots, \quad (1)$$

where the  $\mathbb{Z}_+^p$  valued random vectors

$$\boldsymbol{\xi}_i(n, k), \boldsymbol{\eta}(n), \quad i = 1, \dots, p, \quad n, k = 1, 2, \dots \quad (2)$$

are independent of each other and of the initial value  $\mathbf{X}_0$ , the offspring variables  $\boldsymbol{\xi}_i(n, k)$ ,  $n, k = 1, 2, \dots$ , are identically distributed for every  $i = 1, \dots, p$ , and the innovation variables  $\boldsymbol{\eta}(n)$ ,  $n = 1, 2, \dots$ , are identically distributed as well. Throughout the paper we assume that the offspring variables have finite mean, and we consider the mean matrix

$$\mathbf{M} := E[\boldsymbol{\xi}_1(1, 1), \dots, \boldsymbol{\xi}_p(1, 1)]^\top.$$

It is well-known that the asymptotic behavior of the Galton–Watson process depends largely on the spectral radius  $\varrho(\mathbf{M})$  of the matrix  $\mathbf{M}$ . (See Athreya and Ney (1972), for example.) We say that the process is subcritical, critical or supercritical if the spectral radius is smaller than 1, equal to 1 or larger than 1, respectively. In our paper we investigate only the subcritical (also known as stable) case, that is, we assume that  $\varrho(\mathbf{M}) < 1$ .

Consider an arbitrary state  $\mathbf{x} = (x_1, \dots, x_p)^\top \in \mathbb{Z}_+^p$ . Throughout the paper the notations  $P_{\mathbf{x}}$  and  $E_{\mathbf{x}}$  mean probability and expectation with respect to the condition

$\{\mathbf{X}_0 = \mathbf{x}\}$ . Also, let  $\pi_{\mathbf{x}}^{(n)}(B) = P_{\mathbf{x}}(\mathbf{X}_n \in B)$ ,  $B \subseteq \mathbb{Z}_+^p$ , denotes the distribution of  $\mathbf{X}_n$  under the same initial condition. To shorten the notations we introduce the variable

$$\mathbf{S}(\mathbf{x}) = \sum_{k=1}^{x_1} \boldsymbol{\xi}_1(1, k) + \cdots + \sum_{k=1}^{x_p} \boldsymbol{\xi}_p(1, k), \quad (3)$$

for which we have  $E\mathbf{S}(\mathbf{x}) = \mathbf{M}^\top \mathbf{x}$ . Let  $M_{i,j}^{(n)}$  stands for the  $(i, j)$ -th entry of the matrix  $\mathbf{M}^n$  and let  $\eta_i(1)$  denotes the  $i$ -th component of  $\boldsymbol{\eta}(1)$ . We say that type  $j$  dies out if  $M_{i,j}^{(n)} = 0$  for every  $n \in \mathbb{Z}_+$  and for every type  $i$  such that  $E\eta_i(1) > 0$ , meaning that no type  $j$  being can be born as an offspring of a member of an earlier innovation of any kind. (?)

**Theorem 1.** *If the Galton–Watson process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , is stable then it has an aperiodic communication class  $\mathcal{C}$  such that the process reaches  $\mathcal{C}$  in finitely many steps with probability 1 in case of any initial distribution. Furthermore, if  $\sum_{k=1}^{\infty} \log k P(\|\boldsymbol{\eta}(1)\| = k) < \infty$  then the process has a unique stationary distribution concentrated on the class  $\mathcal{C}$ .*

Let  $\pi$  denotes the unique stationary distribution of the process provided by Theorem 1 and let  $\tilde{\mathbf{X}}$  stands for an arbitrary random vector having distribution  $\pi$ . In several applications showing the linear independence of the components of  $\tilde{\mathbf{X}}$  is required. For example, assume that we want to estimate the mean matrix  $\mathbf{M}$  based on some observations  $\mathbf{X}_0, \dots, \mathbf{X}_n$  by using the conditional least squares method of Klimko and Nelson (1978) or its weighted variant proposed by Wei and Winnicki (1990). Unfortunately, these estimators may not exists for a given realisation of the sample, but using ergodicity one can show that the estimators are well-defined with asymptotic probability 1 if the components of  $\tilde{\mathbf{X}}$  are linearly independent. (For more details on the conditional least squares estimators of higher moments of the offspring and the innovation variables see Nedényi (201?).) Since the stationary distribution  $\pi$  has positive mass at every state  $\mathbf{x} \in \mathcal{C}$  the components of  $\tilde{\mathbf{X}}$  are linearly dependent if and only if the class  $\mathcal{C}$  is a subset of a lower dimensional affine subspace of  $\mathbb{R}^p$ . In our next theorem we provide necessary and sufficient conditions for this behavior.

**Theorem 2.** *Assume that the multitype Galton–Watson process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , is stable. The communication class  $\mathcal{C}$  defined in Theorem 1 is a subset of a lower dimensional affine subspace of  $\mathbb{R}^p$  if and only if either of the following conditions holds:*

- (i) *Any of the types dies out.*
- (ii) *There exists a vector  $\mathbf{c} \in \mathbb{R}^p$ ,  $\mathbf{c} \neq \mathbf{0}$ , such that  $\mathbf{c}^\top \boldsymbol{\xi}_i(1, 1) = 0$  almost surely for  $i = 1, \dots, p$  and the variable  $\mathbf{c}^\top \boldsymbol{\eta}(1)$  is degenerate.*

*Furthermore, if type  $j$  dies out then  $\mathcal{C}$  is a subset of the linear space defined by the equation  $\mathbf{e}_j^\top \mathbf{x} = 0$ ,  $\mathbf{x} \in \mathbb{R}^p$ , and if (ii) is satisfied then  $\mathcal{C}$  is a subset of the affine subpace defined by  $\mathbf{c}^\top \mathbf{x} = \mathbf{c}^\top \boldsymbol{\eta}(1)$ ,  $\mathbf{x} \in \mathbb{R}^p$ .*

We note that condition (i) in Theorem 2 can not be relaxed because it does not imply condition (ii). To prove this let us consider the 2-type Galton–Watson process corresponding to the offspring distributions  $\xi_1$  and  $\xi_2$  and innovation  $\eta$  where the components  $\xi_{2,1}$ ,  $\xi_{2,2}$  and  $\eta_1$  are independent and non-degenerate and  $\xi_{1,2} = \eta_2 = 0$  with probability 1. Then, type 2 dies out but the process does not satisfy condition (ii).

For any given distribution  $\nu$  on  $\mathbb{Z}_+^p$  introduce the norm

$$\|\nu\|_{\mathcal{F}_r} := \sup_{f \in \mathcal{F}_r} \int_{\mathbb{Z}_+^p} f(\mathbf{x}) \nu(d\mathbf{x}),$$

where

$$\mathcal{F}_r = \{f : \mathbb{Z}_+^p \rightarrow \mathbb{R} : |f(\mathbf{x})| \leq \|\mathbf{x}\|^r + 1, \mathbf{x} \in \mathbb{Z}_+^p\}.$$

**Theorem 3.** *Consider a positive integer  $r$ . If the Galton–Watson process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , is stable then the following are equivalent:*

(i) *The stationary distribution  $\pi$  exists and has finite  $r$ -th moment.*

(ii) *We have  $E\|\eta(1)\|^r < \infty$  and  $E\|\xi_i(1,1)\|^r < \infty$  for all types  $i$  that does not die out.*

Furthermore, if (i) or (ii) is satisfied then there exist constants  $a_1 > 1$  and  $a_2 > 0$  such that

$$\sum_{n=0}^{\infty} a_1^n \|\pi_{\mathbf{x}}^{(n)} - \pi\|_{\mathcal{F}_r} \leq a_2 (\|\mathbf{x}\|^r + 1), \quad \mathbf{x} \in \mathcal{C}.$$

Since the indicator function  $\mathbb{1}_B(\mathbf{x})$  of an arbitrary set  $B \subseteq \mathbb{Z}_+^p$  and the function  $f(\mathbf{x}) = \|\mathbf{x}\|^r$  are elements of  $\mathcal{F}_r$  for every  $r$  the following statement is implied by Theorem 3.

**Corollary 4.** *Assume that either (i) or (ii) of Theorem 3 holds with some positive integer  $r$  and consider an arbitrary state  $\mathbf{x} \in \mathcal{C}$ . Then we have the rates*

$$\sup_{B \subseteq \mathbb{Z}_+^p} |P_{\mathbf{x}}(\mathbf{X}_n \in B) - \pi(B)| = o(a_1^n), \quad \sup_{\alpha \in [0, r]} |E_{\mathbf{x}}\|\mathbf{X}_n\|^\alpha - E\|\tilde{\mathbf{X}}\|^\alpha| = o(a_1^n), \quad n \rightarrow \infty.$$

### 3 Proofs

In this section we present the proofs of the results stated in Section 2. The first and the second statement are two lemmas which may be not new results but we provide proofs for them because they are fundamental observation in the subject of our paper.

**Proposition 5.** *Let  $\mathbf{A} \in \mathbb{R}^{p \times p}$  be a matrix having only non-negative entries. If  $\rho(\mathbf{A}) < 1$  then there exist a constant  $\lambda \in (0, 1)$  and a vector  $\mathbf{v} \in \mathbb{R}^p$  such that all components of  $\mathbf{v}$  are strictly positive and  $\mathbf{A}\mathbf{v} \leq \lambda\mathbf{v}$ .*

*Proof.* Because the eigenvalues are continuous functions of the matrix entries we can find  $\varepsilon > 0$  such that  $\lambda := \rho(\mathbf{A} + \varepsilon) < 1$ . Since the matrix  $\mathbf{A} + \varepsilon$  is positive the Perron–Frobenius theorem implies that  $\mathbf{A} + \varepsilon$  has eigenvector  $\mathbf{v}$  with eigenvalue  $\lambda$  such that all components of  $\mathbf{v}$  are strictly positive. With this vector we get the inequality  $\mathbf{A}\mathbf{v} \leq (\mathbf{A} + \varepsilon)\mathbf{v} = \lambda\mathbf{v}$ .  $\square$

**Proposition 6.** *Consider irreducible time-homogeneous  $\mathbb{Z}_+^p$  valued Markov chains  $\mathbf{X}_n$  and  $\mathbf{Y}_n$ ,  $n \in \mathbb{Z}_+$ . If under some initial distribution  $(\mathbf{X}_0, \mathbf{Y}_0)$  we have  $\mathbf{X}_n \leq \mathbf{Y}_n$  almost surely for every  $n \in \mathbb{Z}_+$  then the following hold:*

(i) *If  $\mathbf{Y}_n$ ,  $n \in \mathbb{Z}_+$ , is recurrent then  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , is recurrent.*

(ii) *If  $\mathbf{Y}_n$ ,  $n \in \mathbb{Z}_+$ , is positive recurrent then  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , is positive recurrent.*

*Proof.* Consider any states  $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{Z}_+^p$  such that the event  $A := \{\mathbf{X}_0 = \mathbf{x}_0, \mathbf{Y}_0 = \mathbf{y}_0\}$  has positive probability. Since under this event the inequality  $\mathbf{X}_n \leq \mathbf{Y}_n$  still holds almost surely for every  $n$ , we can condition on  $A$  and we can assume without the loss of generality that the processes have deterministic initial states  $\mathbf{x}_0$  and  $\mathbf{y}_0$ , respectively.

Let  $\mathcal{C} \subseteq \mathbb{Z}_+^p$  denotes the state space of the process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , and introduce the set  $\mathcal{C}_{\mathbf{y}_0} := \{\mathbf{x} \in \mathcal{C} : \mathbf{x} \leq \mathbf{y}_0\}$ . Also, let  $p_{\mathbf{X}}^{(n)}(\cdot, \cdot)$  and  $p_{\mathbf{Y}}^{(n)}(\cdot, \cdot)$  denote the  $n$ -step transition probabilities of the processes. In our proof we will use the well-know characterisation of the types of states by the asymptotic behavior of the transition probabilities, for reference see the main theorem in Section XV.5 of Feller (1968), for example.

Assume that the chain  $\mathbf{Y}_n$ ,  $n \in \mathbb{Z}_+$ , is recurrent. Since the initial distributions are deterministic by assumption we get that

$$\sum_{\mathbf{x} \in \mathcal{C}_{\mathbf{y}_0}} \sum_{n=0}^{\infty} p_{\mathbf{X}}^{(n)}(\mathbf{x}_0, \mathbf{x}) = \sum_{n=0}^{\infty} P(\mathbf{X}_n \in \mathcal{C}_{\mathbf{y}_0}) \geq \sum_{n=0}^{\infty} p_{\mathbf{Y}}^{(n)}(\mathbf{y}_0, \mathbf{y}_0) = \infty,$$

and hence, there exists a state  $\mathbf{x}^* \in \mathcal{C}_{\mathbf{y}_0}$  such that  $\sum_{n=0}^{\infty} p_{\mathbf{X}}^{(n)}(\mathbf{x}_0, \mathbf{x}^*) = \infty$ . The irreducibility of the process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , implies that we have  $p_{\mathbf{X}}^{(k)}(\mathbf{x}^*, \mathbf{x}_0) > 0$  for some  $k \in \mathbb{Z}_+$ . This leads to the inequality

$$\sum_{n=0}^{\infty} p_{\mathbf{X}}^{(n+k)}(\mathbf{x}_0, \mathbf{x}_0) \geq \sum_{n=0}^{\infty} p_{\mathbf{X}}^{(n)}(\mathbf{x}_0, \mathbf{x}^*) p_{\mathbf{X}}^{(k)}(\mathbf{x}^*, \mathbf{x}_0) = \infty$$

proving that state  $\mathbf{x}_0$  is recurrent.

Similarly, if we assume that  $\mathbf{Y}_n$ ,  $n \in \mathbb{Z}_+$ , is positive recurrent then

$$\sum_{\mathbf{x} \in \mathcal{C}_{\mathbf{y}_0}} \limsup_{n \rightarrow \infty} p_{\mathbf{X}}^{(n)}(\mathbf{x}_0, \mathbf{x}) \geq \limsup_{n \rightarrow \infty} P(\mathbf{X}_n \in \mathcal{C}_{\mathbf{y}_0}) \geq \limsup_{n \rightarrow \infty} p_{\mathbf{Y}}^{(n)}(\mathbf{y}_0, \mathbf{y}_0) > 0,$$

and hence,  $\limsup_{n \rightarrow \infty} p_{\mathbf{X}}^{(n)}(\mathbf{x}_0, \mathbf{x}^*) > 0$  for some  $\mathbf{x}^* \in \mathcal{C}_{\mathbf{y}_0}$ . Again, if  $k \in \mathbb{Z}_+$  is a constant such that  $p_{\mathbf{X}}^{(k)}(\mathbf{x}^*, \mathbf{x}_0) > 0$  then

$$\limsup_{n \rightarrow \infty} p_{\mathbf{X}}^{(n+k)}(\mathbf{x}_0, \mathbf{x}_0) \geq \limsup_{n \rightarrow \infty} p_{\mathbf{X}}^{(n)}(\mathbf{x}_0, \mathbf{x}^*) p_{\mathbf{X}}^{(k)}(\mathbf{x}^*, \mathbf{x}_0) > 0.$$

This implies the positive recurrence of state  $\mathbf{x}_0$  and the proof is complete.  $\square$

*Proof of Theorem 1.* We will use the representation of the multitype Galton–Watson process  $\mathbf{X}_n$ ,  $n = 0, 1, \dots$ , provided by Section 2.7 of Mode (1971). Let  $\mathbf{Y}_n$  and  $\mathbf{Z}_n(k)$ ,  $k = 1, \dots, n$ , denote the number of the offsprings of the initial population  $\mathbf{X}_0$  and of the innovation population  $\boldsymbol{\eta}(k)$  living in generation  $n$ , respectively. Using standard calculations one can show that the probability generation function of  $\mathbf{X}_n$  is equal to the product of the probability generating functions of the variables  $\mathbf{Y}_n, \mathbf{Z}_n(1), \dots, \mathbf{Z}_n(n)$ . From this we obtain the representation

$$\mathbf{X}_n = \mathbf{Y}_n + \mathbf{Z}_n(1) + \dots + \mathbf{Z}_n(n) =: \mathbf{Y}_n + \mathbf{V}_n, \quad n = 1, 2, \dots, \quad (4)$$

where  $\mathbf{Y}_n, \mathbf{Z}_n(1), \dots, \mathbf{Z}_n(n)$  are independent of each other. Let us note that the sequence  $\mathbf{Y}_n$ ,  $n = 0, 1, \dots$ , is a multitype Galton–Watson process without immigration, and the assumption  $\varrho(\mathbf{M}) < 1$  implies that  $\mathbf{Y}_n$ ,  $n = 0, 1, \dots$ , becomes extinct with probability 1 in case of any initial distribution  $\mathbf{X}_0$ . (We will show this as a step of the current proof.)

For  $n = 1, 2, \dots$  let  $\mathcal{D}_n \subseteq \mathbb{Z}_+^p$  denote the range of the variable  $\mathbf{V}_n$ , that is, the set of all states  $\mathbf{x} \in \mathbb{Z}_+^p$  such that  $P(\mathbf{V}_n = \mathbf{x}) > 0$ . First we show that  $P_{\mathbf{x}}(\mathbf{X}_1 \in \mathcal{D}_{n+1}) = 1$  for any  $n$  and  $\mathbf{x} \in \mathcal{D}_n$ . Assume that there exists  $\mathbf{x}^* \in \mathcal{D}_n$  such that  $P_{\mathbf{x}^*}(\mathbf{X}_1 \notin \mathcal{D}_{n+1}) > 0$ . Since under the condition  $\mathbf{X}_0 = \mathbf{0}$  we have  $\mathbf{X}_n = \mathbf{V}_n$  we get that

$$\begin{aligned} 0 &= P(\mathbf{X}_{n+1} \notin \mathcal{D}_{n+1} \mid \mathbf{X}_0 = \mathbf{0}) = \sum_{\mathbf{x} \in \mathcal{D}_n} P(\mathbf{X}_{n+1} \in \mathcal{D}_{n+1} \mid \mathbf{X}_n = \mathbf{x}) P(\mathbf{X}_n = \mathbf{x} \mid \mathbf{X}_0 = \mathbf{0}) \\ &= \sum_{\mathbf{x} \in \mathcal{D}_n} P_{\mathbf{x}}(\mathbf{X}_1 \in \mathcal{D}_{n+1}) P(\mathbf{V}_n = \mathbf{x}) \geq P_{\mathbf{x}^*}(\mathbf{X}_1 \in \mathcal{D}_{n+1}) P(\mathbf{V}_n = \mathbf{x}^*) > 0, \end{aligned}$$

which is a contradiction. As a consequence the set  $\mathcal{C}_n := \cup_{k=n}^{\infty} \mathcal{D}_k \subseteq \mathbb{Z}_+^p$  is closed for any  $n$  in the sense that  $P_{\mathbf{x}}(\mathbf{X}_1 \in \mathcal{C}_n) = 1$  holds for every  $\mathbf{x} \in \mathcal{C}_n$ .

Let us recall that the sequence  $\mathbf{Y}_n$ ,  $n = 0, 1, \dots$ , is a multitype Galton–Watson process without immigration. As the next step of the proof we show that this process may die out really fast. Let  $\mathbf{v} \in \mathbb{R}_+^p$  and  $\lambda \in (0, 1)$  be the vector and the constant provided by Lemma 5 with  $\mathbf{A} = \mathbf{M}$ . Since  $\mathbf{Y}_n = \mathbf{S}(\mathbf{Y}_{n-1})$ ,  $n = 1, 2, \dots$ , we obtain the inequality

$$0 \leq \mathbf{v}^\top E_{\mathbf{x}}(\mathbf{Y}_n) = \mathbf{v}^\top E_{\mathbf{x}}(E[\mathbf{S}(\mathbf{Y}_{n-1}) \mid \mathbf{Y}_{n-1}]) = \mathbf{v}^\top E_{\mathbf{x}}(\mathbf{M}^\top \mathbf{Y}_{n-1}) \leq \lambda \mathbf{v}^\top E_{\mathbf{x}}(\mathbf{Y}_{n-1}) = \dots \leq \lambda^n \mathbf{v}^\top \mathbf{x}, \blacksquare$$

and hence,  $E_{\mathbf{x}}(\mathbf{Y}_n) \rightarrow 0$  as  $n \rightarrow \infty$  in case of any initial state  $\mathbf{x} \in \mathbb{Z}_+^p$ . Because  $\mathbf{Y}_n$  has non-negative integer components we have  $P_{\mathbf{x}}(\mathbf{Y}_n = \mathbf{0}) \rightarrow 1$ , that is, there exists an integer  $n^*(\mathbf{x})$  such that  $P_{\mathbf{x}}(\mathbf{Y}_n = \mathbf{0}) > 0$  for every  $n \geq n^*(\mathbf{x})$ . Since the number of the offsprings of the members of the initial population  $\mathbf{x}$  are independent of each other, we obtain the inequality

$$P_{\mathbf{x}}(\mathbf{Y}_n = \mathbf{0}) = P_{\mathbf{e}_1}(\mathbf{Y}_n = \mathbf{0})^{x_1} \dots P_{\mathbf{e}_p}(\mathbf{Y}_n = \mathbf{0})^{x_n} > 0$$

for every  $n \geq n^* := \max(n^*(\mathbf{e}_1), \dots, n^*(\mathbf{e}_p))$ . That is, it has a positive probability that the process  $\mathbf{Y}_n$ ,  $n = 0, 1, \dots$ , becomes extinct in  $n^*$  steps in case of any initial value  $\mathbf{x}$ .

Fix an integer  $n \geq n^*$  and consider arbitrary states  $\mathbf{x} \in \mathbb{Z}_+^p$  and  $\mathbf{z} \in \mathcal{D}_n$ . Since the variables  $\mathbf{Y}_n$  and  $\mathbf{V}_n$  are independent of each other in case of any initial state, we get that

$$P_{\mathbf{x}}(\mathbf{X}_n = \mathbf{z}) \geq P_{\mathbf{x}}(\mathbf{Y}_n = \mathbf{0}, \mathbf{Z}_n = \mathbf{z}) = P_{\mathbf{x}}(\mathbf{Y}_n = \mathbf{0}) P(\mathbf{Z}_n = \mathbf{z}) > 0, \quad (5)$$

meaning that every elements of  $\mathcal{D}_n$  are accessible from the arbitrary state  $\mathbf{x}$  in  $n$  steps. As a consequence we get that the elements of the set  $\mathcal{C} := \mathcal{C}_{n^*}$  communicate with each other. Since  $\mathcal{C}$  is closed it is a communication class of the process  $\mathbf{X}_n, n \in \mathbb{Z}_+$ . Consider an arbitrary state  $\mathbf{z} \in \mathcal{D}_n$  and a non-negative integer  $m$  and let  $\mathbf{x} \in \mathbb{Z}_+^p$  be a state such that  $P_{\mathbf{z}}(\mathbf{X}_m = \mathbf{x}) > 0$ . Using equation (5) again we find that

$$P_{\mathbf{z}}(\mathbf{X}_{n+m} = \mathbf{z}) \geq P_{\mathbf{z}}(\mathbf{X}_m = \mathbf{x})P_{\mathbf{x}}(\mathbf{X}_n = \mathbf{z}) > 0,$$

that is, state  $\mathbf{z}$  is accessible from itself in  $n + m$  steps. Since  $m$  was an arbitrary non-negative integer the communication class  $\mathcal{C}$  is aperiodic.

To prove the first statement of the theorem it is only remained to show that the process  $\mathbf{X}_n, n \in \mathbb{Z}_+$ , reaches the class  $\mathcal{C}$  in finitely many steps with probability 1 in case of any initial distribution  $\mathbf{X}_0$ . Since the state space is countable it is enough to prove this statement under the condition  $\{\mathbf{X}_0 = \mathbf{x}\}$  where  $\mathbf{x} \in \mathbb{Z}_+^p$  is an arbitrary fixed state. Using that the events  $\{\mathbf{Y}_n = \mathbf{0}\}, n \in \mathbb{Z}_+$ , form an increasing sequence we get that

$$\begin{aligned} P_{\mathbf{x}}(\exists n \geq n^* : \mathbf{X}_n \in \mathcal{C}) &\geq P_{\mathbf{x}}(\exists n \geq n^* : \mathbf{X}_n \in \mathcal{D}_n) \geq P_{\mathbf{x}}(\exists n \geq n^* : \mathbf{Y}_n = \mathbf{0}) \\ &= P_{\mathbf{x}}(\cup_{n=n^*}^{\infty} \{\mathbf{Y}_n = \mathbf{0}\}) = \lim_{n \rightarrow \infty} P_{\mathbf{x}}(\mathbf{Y}_n = \mathbf{0}) = 1. \end{aligned}$$

For the second statement note again that the eigenvalues depends continuously on the matrix entries. Since our process is stable there exists  $\varepsilon > 0$  such that  $\varrho(\mathbf{M}') < 1$  holds with  $\mathbf{M}' = \mathbf{M} + \varepsilon$ . Let  $\mathbf{1} \in \mathbb{R}^p$  denotes the vector whose every components are equal with 1 and consider independent and identically distributed random vectors  $\mathbb{1}_i(n, k)$ ,  $i = 1, \dots, p, n, k = 1, 2, \dots$ , being independent of the variables in formula (2) and having common distribution  $P(\mathbb{1}_i(n, k) = \mathbf{1}) = \varepsilon$  and  $P(\mathbb{1}_i(n, k) = \mathbf{0}) = 1 - \varepsilon$ . Consider the multitype Galton–Watson process  $\mathbf{X}'_n, n \in \mathbb{Z}_+$ , obtained by replacing the offspring variables  $\xi_i(n, k)$  in (1) with

$$\xi'_i(n, k) := \xi_i(n, k) + \mathbb{1}_i(n, k), \quad i = 1, \dots, p, \quad n, k = 1, 2, \dots, \quad (6)$$

and let  $\mathcal{C}'$  denote the unique closed communication class of the processes  $\mathbf{X}'_n, n \in \mathbb{Z}_+$ , provided by the first statement of the current theorem. It follows from the construction that the set  $\mathcal{C}'$  is not bounded and the mean matrix of the new offspring variables is  $E[\xi'_1(1, 1), \dots, \xi'_p(1, 1)]^\top = \mathbf{M}'$ . Since the matrix  $\mathbf{M}'$  is positive and  $\varrho(\mathbf{M}') < 1$ , Corollary 1 of Kaplan (1973) implies that the process  $\mathbf{X}'_n, n \in \mathbb{Z}_+$ , has a stationary distribution which must be concentrated on  $\mathcal{C}'$ . Consider any states  $\mathbf{x}_0 \in \mathcal{C}$  and  $\mathbf{x}'_0 \in \mathcal{C}'$  such that  $\mathbf{x}_0 \leq \mathbf{x}'_0$ . Under the event  $\{\mathbf{X}_0 = \mathbf{x}_0, \mathbf{X}'_0 = \mathbf{x}'_0\}$  both chains are irreducible and it follows by induction that we have  $\mathbf{X}_n \leq \mathbf{X}'_n$  almost surely for every  $n$ . Since the process  $\mathbf{X}'_n, n \in \mathbb{Z}_+$ , is positive recurrent statement (ii) of our Lemma 6 implies that  $\mathbf{X}_n, n \in \mathbb{Z}_+$ , is positive recurrent too, completing the proof.  $\square$

*Proof of Theorem 2.* First we show that if either (i) or (ii) is satisfied then  $\mathcal{C}$  is a subset of a lower dimensional affine subspace  $\mathcal{S}$  of  $\mathbb{R}^p$ . Assume that type  $j$  dies out for some  $j = 1, \dots, p$ . Then, by using the notations introduced in the proof of Theorem 1, the  $j$ -th component of  $\mathbf{V}_n$  vanishes with probability 1 for every positive integer  $n$ . Since  $\mathcal{C}$  is

defined as the union of the ranges of the variables  $\mathbf{V}_n$ ,  $n \geq n^*$ , the class  $\mathcal{C}$  is a subset of the linear subspace  $\mathcal{S}$  defined by the equation  $\mathbf{e}_j^\top \mathbf{v} = 0$ ,  $\mathbf{v} \in \mathbb{R}^p$ .

Now assume that (ii) holds and consider an arbitrary  $\mathbf{x}' \in \mathcal{C}$ . Since  $\mathcal{C}$  is a communication class the state  $\mathbf{x}'$  is accesible from some  $\mathbf{x} \in \mathcal{C}$  in one step. Working on the event  $\{\mathbf{X}_0 = \mathbf{x}\}$  we get the almost sure equation

$$\mathbf{c}^\top \mathbf{X}_1 = \sum_{i=1}^p \sum_{k=1}^{x_i} \mathbf{c}^\top \boldsymbol{\xi}_i(1, k) + \mathbf{c}^\top \boldsymbol{\eta}(1) = 0 + \mathbf{c}^\top \boldsymbol{\eta}(1),$$

where  $\mathbf{c}^\top \boldsymbol{\eta}(1)$  is degenerate by assumption. Since  $P_{\mathbf{x}}(\mathbf{X}_1 = \mathbf{x}') > 0$  we get that  $\mathbf{x}'$  is an element of the affine subspace  $\mathcal{S}$  defined by the equation  $\mathbf{c}^\top \mathbf{v} = \mathbf{c}^\top \boldsymbol{\eta}(1)$ ,  $\mathbf{v} \in \mathbb{R}^p$ .

For the contrary direction let  $\mathcal{S} \subsetneq \mathbb{R}^p$  denotes the affine subspace generated by  $\mathcal{C}$  and assume that none of the types dies out. Consider an arbitrary state  $\mathbf{x}^* \in \mathcal{C}$  and fix a vector  $\mathbf{y}^* \in \mathbb{Z}_+^p$  such that  $P(\boldsymbol{\eta}(1) = \mathbf{y}^*) > 0$ . Since the set  $\mathcal{V} := \mathcal{S} - \mathbf{x}^*$  is a linear subspace of  $\mathbb{R}^p$  with dimension less than  $p$ , the orthogonal complement  $\mathcal{V}^\perp$  of  $\mathcal{V}$  is a non-trivial linear subspace of  $\mathbb{R}^p$  and we have  $\mathbf{c}^\top (\mathbf{x} - \mathbf{x}^*) = 0$  for every  $\mathbf{c} \in \mathcal{V}^\perp$  and  $\mathbf{x} \in \mathcal{S}$ .

Consider an arbitrary state  $\mathbf{x} \in \mathcal{C}$  and an arbitrary vector  $\mathbf{c} \in \mathcal{V}^\perp$  and work on the event  $\{\mathbf{X}_0 = \mathbf{x}\}$ . Since the communication class  $\mathcal{C}$  is closed the variable  $\mathbf{X}_1$  lies in  $\mathcal{S}$  with probability 1 and we get the almost sure equation

$$\mathbf{c}^\top \mathbf{x}^* = \mathbf{c}^\top \mathbf{X}_1 = \sum_{k=1}^{x_1} \mathbf{c}^\top \boldsymbol{\xi}_1(1, k) + \dots + \sum_{k=1}^{x_p} \mathbf{c}^\top \boldsymbol{\xi}_p(1, k) + \mathbf{c}^\top \boldsymbol{\eta}(1). \quad (7)$$

Since the left side of this equation is deterministic and the terms on the right side are independent of each other we get that these terms are deterministic as well. Consider any type  $i = 1, \dots, p$ . Because we assumed that type  $i$  does not die out the state  $\mathbf{x} \in \mathcal{C}$  in (7) can be chosen such that  $x_i \neq 0$ . This implies that  $\mathbf{c}^\top \boldsymbol{\xi}_i(1, 1)$  and  $\mathbf{c}^\top \boldsymbol{\eta}(1)$  are degenerate variables,  $i = 1, \dots, p$ , and by using formula (3) we get that the equation

$$\mathbf{c}^\top \mathbf{x}^* = \mathbf{c}^\top \mathbf{X}_1 = \mathbf{c}^\top \mathbf{S}(\mathbf{x}) + \mathbf{c}^\top \boldsymbol{\eta}(1) = E_{\mathbf{x}}(\mathbf{c}^\top \mathbf{S}(\mathbf{x})) + \mathbf{c}^\top \mathbf{y}^* = (\mathbf{M}\mathbf{c})^\top \mathbf{x} + \mathbf{c}^\top \mathbf{y}^* \quad (8)$$

holds almost surely in case of any state  $\mathbf{x} \in \mathcal{C}$  and vector  $\mathbf{c} \in \mathcal{V}^\perp$ . Since  $\mathcal{S}$  is the affine subspace generated by the set  $\mathcal{C}$ , equation (8) is valid for any  $\mathbf{x} \in \mathcal{S}$ , as well.

Consider an arbitrary vector  $\mathbf{v} \in \mathcal{V}$ . Since both  $\mathbf{v} + \mathbf{x}^*$  and  $\mathbf{x}^*$  are elements of  $\mathcal{S}$ , from equation (8) it follows that

$$(\mathbf{M}\mathbf{c})^\top \mathbf{v} = (\mathbf{M}\mathbf{c})^\top (\mathbf{v} + \mathbf{x}^*) - (\mathbf{M}\mathbf{c})^\top \mathbf{x}^* = \mathbf{c}^\top (\mathbf{x}^* - \mathbf{y}^*) - \mathbf{c}^\top (\mathbf{x}^* - \mathbf{y}^*) = 0.$$

This implies that  $\mathbf{M}\mathbf{c} \in \mathcal{V}^\perp$  for any  $\mathbf{c} \in \mathcal{V}^\perp$ , and hence, by introducing the linear function

$$\psi : \mathbb{R}^p \rightarrow \mathbb{R}^p, \quad \psi(\mathbf{c}) = \mathbf{M}\mathbf{c},$$

we have that  $\psi(\mathcal{V}^\perp) \subseteq \mathcal{V}^\perp$ . Because the process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , reaches the class  $\mathcal{C}$  almost surely in finitely many steps in case of any initial state, there exists a state  $\mathbf{z} \in \mathbb{Z}_+^p$ ,  $\mathbf{z} \notin \mathcal{S}$ ,

such that the subspace  $\mathcal{S}$  is accesible from  $\mathbf{z}$  in one step. Since under the event  $\{\mathbf{X}_0 = \mathbf{z}\}$  the variable  $\mathbf{X}_1$  is an element of  $\mathcal{S}$  with positive probability, the equations

$$\mathbf{c}^\top \mathbf{x}^* = \mathbf{c}^\top \mathbf{X}_1 = \mathbf{c}^\top \mathbf{S}(\mathbf{z}) + \mathbf{c}^\top \boldsymbol{\eta}(1) = E_{\mathbf{x}}(\mathbf{c}^\top \mathbf{S}(\mathbf{z})) + \mathbf{c}^\top \mathbf{y}^* = (\mathbf{M}\mathbf{c})^\top \mathbf{z} + \mathbf{c}^\top \mathbf{y}^* \quad (9)$$

hold with positive probability in case of any  $\mathbf{c} \in \mathcal{V}^\perp$ . As a consequence, we get that  $\mathbf{c}^\top (\mathbf{x}^* - \mathbf{y}^*) = (\mathbf{M}\mathbf{c})^\top \mathbf{z}$ . Let us consider the orthogonal decomposition  $\mathbf{z} = \mathbf{x} + \mathbf{x}^\perp$  where  $\mathbf{x} \in \mathcal{S}$  and  $\mathbf{x}^\perp \in \mathcal{V}^\perp$ ,  $\mathbf{x}^\perp \neq \mathbf{0}$ . From (8) we obtain that

$$\mathbf{c}^\top (\mathbf{x}^* - \mathbf{y}^*) = (\mathbf{M}\mathbf{c})^\top \mathbf{z} = (\mathbf{M}\mathbf{c})^\top \mathbf{x} + (\mathbf{M}\mathbf{c})^\top \mathbf{x}^\perp = \mathbf{c}^\top (\mathbf{x}^* - \mathbf{y}^*) + \psi(\mathbf{c})^\top \mathbf{x}^\perp,$$

and hence  $\psi(\mathbf{c})^\top \mathbf{x}^\perp = 0$  for any vector  $\mathbf{c} \in \mathcal{V}^\perp$ . That is,  $\psi(\mathcal{V}^\perp) \perp \mathbf{x}^\perp \in \mathcal{V}^\perp$  implying that  $\psi(\mathcal{V}^\perp) \subsetneq \mathcal{V}^\perp$ . This means that  $\psi$  is not a full rank linear transformation and there exists a vector  $\mathbf{c}^* \in \mathcal{V}^\perp$  such that  $\mathbf{M}\mathbf{c}^* = \psi(\mathbf{c}^*) = \mathbf{0}$ . Since the variables  $\mathbf{c}^\top \boldsymbol{\xi}_i(1, 1)$ ,  $i = 1, \dots, p$ , are deterministic in case of any  $\mathbf{c} \in \mathcal{V}^\perp$ , we get that

$$(\mathbf{c}^*)^\top \boldsymbol{\xi}_i(1, 1) = E((\mathbf{c}^*)^\top \boldsymbol{\xi}_i(1, 1)) = (\mathbf{c}^*)^\top E\boldsymbol{\xi}_i(1, 1) = (\mathbf{M}\mathbf{c}^*)^\top \mathbf{e}_i = 0, \quad i = 1, \dots, p,$$

and the proof is complete.  $\square$

Our next result is a slight generalisation of Lemma 9.2 of Barczy et al. (201?)

**Proposition 7.** *Consider any positive integer  $r$ . There exists a polynomial  $Q_r$  of degree at most  $\lfloor r/2 \rfloor$  such that for any independent zero-mean random variables  $\zeta_1, \dots, \zeta_\ell$  we have the inequality*

$$|E(\zeta_1 + \dots + \zeta_\ell)^r| \leq b_r Q_r(\ell),$$

where  $b_r = \sup_{1 \leq i \leq \ell} E|\zeta_i|^r$ .

*Proof.* Consider the constants  $b_q := \sup_{1 \leq i \leq \ell} E|\zeta_i|^q \leq b_r^{q/r}$ ,  $q = 1, \dots, r$ . Since the variables are zero-mean we have  $E(\zeta_1^{r_1} \dots \zeta_\ell^{r_\ell}) = 0$  if any of the exponents is equal with 1. From this we get that

$$\begin{aligned} |E(\zeta_1 + \dots + \zeta_\ell)^r| &\leq \sum_{\substack{r_1, \dots, r_\ell \in \mathbb{Z}_+ \\ r_1 + \dots + r_\ell = r}} \frac{r!}{r_1! \dots r_\ell!} |E(\zeta_1^{r_1} \dots \zeta_\ell^{r_\ell})| \leq \sum_{\substack{r_1, \dots, r_\ell \in \mathbb{Z}_+ \setminus \{0\} \\ r_1 + \dots + r_\ell = r}} \frac{r!}{r_1! \dots r_\ell!} b_{r_1} \dots b_{r_\ell} \\ &\leq \sum_{\substack{k_2, \dots, k_r \in \mathbb{Z}_+ \\ 2k_2 + \dots + rk_r = r}} \binom{\ell}{k_2} \binom{\ell - k_2}{k_3} \dots \binom{\ell - (k_2 + \dots + k_{r-1})}{k_r} \frac{r!}{(2!)^{k_2} \dots (r!)^{k_r}} b_2^{k_2} \dots b_r^{k_r} \\ &\leq b_r \sum_{\substack{k_2, \dots, k_r \in \mathbb{Z}_+ \\ 2k_2 + \dots + rk_r = r \\ k_2 + \dots + k_r \leq \ell}} \frac{\ell(\ell - k_2) \dots (\ell - k_2 - \dots - k_r + 1)}{k_2! \dots k_r!} \frac{r!}{(2!)^{k_2} \dots (r!)^{k_r}} =: b_r Q_r(\ell). \end{aligned}$$

Since  $2(k_2 + \dots + k_r) \leq r$  every terms of the polynomial  $Q_r$  are of degree at most  $\lfloor r/2 \rfloor$ .  $\square$



*Proof of Theorem 3.* First assume that the unique stationary distribution  $\pi$  provided by Theorem 1 has finite  $r$ -th moment and consider an arbitrary type  $j$  that does not die out. Then there exists a state  $\mathbf{x}^* \in \mathcal{C}$  whose  $j$ -th component is not zero. (To see this start the process from an arbitrary fixed  $\mathbf{x} \in \mathcal{C}$ . Since type  $j$  does not die out the process can reach such a state  $\mathbf{x}^*$  in finitely many steps with positive probability.) By starting the process with the stationary distribution the variable  $\mathbf{X}_1$  has  $r$ -th moment

$$\infty > E\|\mathbf{X}_1\|^r \geq E_{\mathbf{x}^*}\|\mathbf{X}_1\|^r P(\mathbf{X}_0 = \mathbf{x}^*) \geq E\|\boldsymbol{\xi}_j(1, 1)\|^r \pi(\{\mathbf{x}^*\})$$

proving that  $E\|\boldsymbol{\xi}_j(1, 1)\|^r$  is finite.

For the contrary direction assume that (ii) holds with some fixed positive integer  $r$ . By Theorem 2 if type  $j$  dies out then the  $j$ -th components of the states in  $\mathcal{C}$  are zero. Since the invariant distribution  $\pi$  provided by Theorem 1 is concentrated on  $\mathcal{C}$  and the vector norm is defined as the sum of the absolute values of the components, the moments of the stationary distribution do not change if we reduce the dimension of the problem by omitting all types that die out. That is, we can assume without the loss of generality that none of the types dies out.

Consider the constant  $\lambda \in (0, 1)$  and the vector  $\mathbf{v} \in \mathbb{Z}_+^p$  of Lemma 5 with  $\mathbf{A} = \mathbf{M}$ . Since the lemma remains true if we multiply  $\mathbf{v}$  with an arbitrary positive value we can assume that all components of  $\mathbf{v}$  are larger than 1. Also, introduce the function  $V(\mathbf{x}) = (\mathbf{v}^\top \mathbf{x})^r + 1$ ,  $\mathbf{x} \in \mathbb{Z}_+^p$ . Our goal is to prove that

$$E_{\mathbf{x}}V(\mathbf{X}_1) - V(\mathbf{x}) \leq -\beta V(\mathbf{x}) + \gamma \mathbb{1}_{\mathcal{C}'}(\mathbf{x}), \quad \mathbf{x} \in \mathcal{C}, \quad (10)$$

holds where  $\beta > 0$  and  $\gamma < \infty$  are suitable constants and  $\mathbb{1}_{\mathcal{C}'}$  is the indicator function of a suitable finite set  $\mathcal{C}' \subseteq \mathcal{C}$ . Since under the condition  $\{\mathbf{X}_0 = \mathbf{x}\}$  the process  $\mathbf{X}_n$ ,  $n \in \mathbb{Z}_+$ , is irreducible, the finite set  $\mathcal{C}'$  is *small*, and hence, it is *petite* in the sense of Meyn and Tweedie (2009). (See the definition of small sets and Proposition 5.5.3 in their book.) Also note that the class  $\mathcal{C}$  is aperiodic and we have  $\|\mathbf{x}\|^r + 1 \leq V(\mathbf{x})$  for every  $\mathbf{x} \in \mathcal{C}$ . Then, if we can prove the Foster–Lyapunov criteria in (10) then Theorem 15.0.1 of Meyn and Tweedie (2009) immediately implies (i) and the additional statement of the current theorem.

Let  $\mathbf{x} \in \mathcal{C}$  be an arbitrary initial value and introduce the variables  $\overline{\mathbf{X}}_1 := \mathbf{X}_1 - E_{\mathbf{x}}\mathbf{X}_1$  and

$$\overline{\boldsymbol{\xi}}_i(n, k) := \boldsymbol{\xi}_i(n, k) - E\boldsymbol{\xi}_i(n, k), \quad \overline{\boldsymbol{\eta}}(n) := \boldsymbol{\eta}(n) - E\boldsymbol{\eta}(n), \quad i = 1, \dots, p, \quad n, k = 1, 2, \dots$$

Also, note that the moment condition on the offspring and the innovation distributions implies that

$$b_q := \max \left\{ E|\mathbf{v}^\top \overline{\boldsymbol{\xi}}_1(1, 1)|^q, \dots, E|\mathbf{v}^\top \overline{\boldsymbol{\xi}}_p(1, 1)|^q, E|\mathbf{v}^\top \overline{\boldsymbol{\eta}}(1)|^q \right\} < \infty, \quad q = 1, \dots, r.$$

Then, with the polynomial  $Q_q$  defined in Proposition 7 we have

$$|E_{\mathbf{x}}(\mathbf{v}^\top \overline{\mathbf{X}}_1)^q| = \left| E \left( \sum_{k=1}^{x_1} \mathbf{v}^\top \overline{\boldsymbol{\xi}}_1(n, k) + \dots + \sum_{k=1}^{x_p} \mathbf{v}^\top \overline{\boldsymbol{\xi}}_p(n, k) + \mathbf{v}^\top \overline{\boldsymbol{\eta}}(n) \right)^q \right| \leq B_q Q_q(\|\mathbf{x}\| + 1).$$

From this we obtain that

$$\begin{aligned}
E_{\mathbf{x}}V(\mathbf{X}_1) &= E_{\mathbf{x}}\left(\mathbf{v}^\top(\overline{\mathbf{X}}_1 + E_{\mathbf{x}}\mathbf{X}_1)\right)^r + 1 = E_{\mathbf{x}}\left(\mathbf{v}^\top\overline{\mathbf{X}}_1 + (\mathbf{M}\mathbf{v})^\top\mathbf{x} + \mathbf{v}^\top E\boldsymbol{\eta}(1)\right)^r + 1 \\
&\leq \sum_{\substack{q,i,j \in \mathbb{Z}_+ \\ i+j+q=r}} \frac{r!}{q!i!j!} \left|E_{\mathbf{x}}(\mathbf{v}^\top\overline{\mathbf{X}}_1)^q\right| (\lambda\mathbf{v}^\top\mathbf{x})^i (\mathbf{v}^\top E\boldsymbol{\eta}(1))^j + 1 \\
&\leq (\lambda\mathbf{v}^\top\mathbf{x})^r + \sum_{\substack{q,i,j \in \mathbb{Z}_+, i \neq r \\ q+i+j=r}} \frac{r!}{q!i!j!} b_q Q_q(\|\mathbf{x}\| + 1) (\lambda\|\mathbf{v}\|\|\mathbf{x}\|)^i (\mathbf{v}^\top E\boldsymbol{\eta}(1))^j + 1 \\
&=: \lambda^r V(\mathbf{x}) + (1 - \lambda^r) + R(\|\mathbf{x}\|),
\end{aligned}$$

where  $R$  is a polynomial of degree at most  $r - 1$ . As a result it follows that

$$E_{\mathbf{x}}V(\mathbf{X}_1) - V(\mathbf{x}) = -2\beta V(\mathbf{x}) + 2\beta + R(\|\mathbf{x}\|), \quad \mathbf{x} \in \mathcal{C}, \quad (11)$$

where  $\beta = (1 - \lambda^r)/2$ .

Consider the set  $\mathcal{C}' := \{\mathbf{x} \in \mathcal{C} : 2\beta + R(\|\mathbf{x}\|) > \beta V(\mathbf{x})\}$ . Since all components of  $\mathbf{v}$  are larger than 1 we have  $V(\mathbf{x}) \geq \|\mathbf{x}\|^r$ ,  $\mathbf{x} \in \mathcal{C}$ , which implies that  $\mathcal{C}'$  is finite. Then

$$\gamma := \max_{\mathbf{x} \in \mathcal{C}'} \{2\beta + R(\|\mathbf{x}\|) - \beta V(\mathbf{x})\}$$

is a finite real value and from formula (11) we immediately obtain the desired inequality (10). This completes the proof.  $\square$

## References

- Athreya, K.B. and Ney, P. (1972) Branching Processes. Springer, Berlin-Heidelberg-New York.
- Barczy, M., Ispány, M. and Pap, Gy. (201?) Asymptotic behavior of conditional least squares estimators for unstable integer-valued autoregressive models of order 2. Scand. J. Stat., accepted.
- Feller, W. (1968) An introduction to probability theory and its applications. Vol. I. John Wiley & Sons, Inc., New York-London-Sydney.
- Kaplan, N. (1973) The multitype Galton-Watson process with immigration. Ann. Probab. **1** 947-953.
- Klimko, L.A. and Nelson, P.I. (1978) On conditional least squares estimation for stochastic processes. Ann. Stat. **6** 629-642.
- Meyn, S.P. and Tweedie, R.L. (2009) Markov Chains and Stochastic Stability. Cambridge University Press, Cambridge.

- Mode, C.J. (1971) Multitype Branching Processes, Theory and Applications. American Elsevier Publishing Co., Inc., New York.
- Nedényi, F. (201?) Online change-point detection in multitype Galton–Watson processes. manuscript.
- Wei, C.Z. and Winnicki, J. (1990) Estimation of the means in the branching process with immigration. *Ann. Stat.* **18** 1757-1773.